Original Article

Prognostic factors for the survival of patients with esophageal cancer in Northern Iran

Mahmood Reza Ghadimi¹, Mahboobeh Rasouli¹, Mahmood Mahmoodi², Kazem Mohammad³

Abstract

BACKGROUND: Esophageal cancer is the 8th most common cancer and the 6th leading cause of cancer-related death, worldwide. In Iran, the high incidence rates of this type of cancer have been reported from the Caspian Sea region. This study aimed at assessing the factors affecting survival of patients with esophageal cancer in neighbor provinces around Caspian Sea using parametric and semi-parametric models with univariate gamma frailty model.

METHODS: In this study, we performed a prospective review of 359 patients presenting with esophageal cancer from 1990 to 1991. The data were obtained using the Cancer Registry information existed in Babol research center in Iran. Study participants were followed-up until 2006 for a period of 15 years. Hazard ratio was used to interpret the risk of death. The Akaike Information Criterion (AIC) was considered as a criterion to select the best model(s).

RESULTS: Of the 359 patients, 225 (62.7%) were male with a mean age of 60.0 years and 134 (37.3%) were female with a mean age of 55.3 at the time of diagnosis. 1-, 3- and 5-year survival rates after diagnosis were 23%, 15% and 13% , respectively. Comparison between Cox and parametric models of AIC showed that the overall fitting was improved under parametric models. Among parametric models, the log-logistic model with gamma frailty provided better performance than other models. Using this model, we found that gender (p=0.012) and family history of cancer (p= 0.003) were significant predictors.

CONCLUSIONS: Since the proportionality assumption of the Cox model was not held (p = 0.01), the Cox regression model was not an appropriate choice for analyzing our data. According to our findings, log logistic model with gamma frailty could be considered as a useful statistical model in survival analysis of patients with esophageal cancer rather than Cox and log-normal models.

KEYWORDS: Esophageal Cancer; Survival Analysis; Univariate Gamma Frailty Model; Parametric Model.
doubt that the diagnosis of life-threatening diseases such as cancer makes different effects on the patients’ quality of life. Geographical distribution is effective in esophageal cancer. Approximately 80% of the total cancer cases occur in developing countries. Esophageal cancer in the Western countries is relatively rare, but China, South Africa and North of Central Asia regions are considered as having the highest incidence of esophageal cancer. The northern regions of Iran have been observed to dominate malignancies of esophageal cancer. The highest incidence of esophageal cancer occurs in people aged 50-70 years, also the frequency of the disease is higher in men. Overall, gastrointestinal (GI) cancers account for approximately half (44.4%) of all cancer-related deaths in Iran. Unfortunately, patients suffering from esophageal cancer often refer to medical care when it is at advanced stages and so limited or no effective therapies are available to treat them. Theoretically, esophageal cancer cases may be treatable in their early stages, therefore, early detection is desirable.

Cox proportional hazards model is among the most common models in survival data analysis. In cases where the premise of proportional hazards is not met, however, this model will not be appropriate to use. One of the cases in which the Cox proportional hazards model should not be used is when the hazards increase initially but start to decrease after a period of time. In such cases, utilizing log-logistic and log-normal parametric models seem proper, due to having the non-monotonic hazards functions. In the present study, because of facing such models for the hazards function, the log-logistic and log-normal parametric models are used.

Based on asymptotic results, Efron and Oakes showed that under certain circumstances, parameter estimates in parametric models are led to become more efficient than Cox’s model. Selected parametric models such as Weibull, log-logistic and log-normal models are alternatives to the Cox model.

In the Cox proportional hazard model and parametric models, it is assumed that if individuals have the same values of the covariates, they have the same survival function regardless of heterogeneities. But, extra-heterogeneities might exist, that we do not include in the model. We would collect data on all factors we think will influence survival, but there will always be others which have an additional effect that we miss.

A model that is becoming increasingly popular for modeling association between individual survival times within subgroups is the use of a frailty model. A frailty is an unobservable random effect shared by subjects within a subgroup. Frailty models are also used in making adjustments for overdispersion in univariate survival studies. Here, the frailty represents the total effect on survival of the covariates not measured when collecting information on individual subjects. If these effects are ignored, the resulting survival estimates may be misleading. Corrections for this overdispersion allow for adjustments for other unmeasured important effects. The overdispersion in this case is indicated by an unobservable multiplicative effect on the hazard, or frailty. Since the hazard function cannot be negative, a positive distribution should be considered for frailty distribution. The frailty distributions most often applied are the gamma distribution, inverse Gaussian, log-normal, the positive stable distribution, Compound Poisson and a three-parameter distribution [Power Variance Function (PVF)]. Also, because of over-parameterization and identifiability problem, and as frailty is considered as a random effect that indicates the effect of unknown variables, it is necessary to assume that the mean of frailty equals one. The aim of the present study was to assess the factors influencing the survival of patients with esophageal cancer using parametric models with univariate gamma frailty model and to compare the results with those obtained in Cox model.
Methods
This survey was a prospective study. 359 patients with developed esophageal cancer registered at the Babol Cancer Registration Center during 1990-1991 were included in the study. They, then, were followed up for 15 years until 2006. The socio-demographic and clinical data were obtained using questionnaire and the patients’ clinical records. The factors we considered in our study were age at diagnosis, gender, place of residence, province, type of cancer, method of cancer detection, family history of cancer, education, job, marital status, cigarette smoking, ethnicity, migration status, and drug use.

The coding of the samples was done under the direct supervision of pathology specialists based on the international classification of disease for oncology (ICD-O) coding. The study was confirmed by the Ethics Committee of Tehran University of Medical Sciences. In this study, in order to compare the efficiency of parametric and Cox models the AIC (Akaike Information Criterion) was used. The AIC is a criterion that assesses goodness of fit of a statistical model, and the lower value of AIC suggests a better model.

In multiple analysis, RR (relative risk) and HR (hazard rate) were used to interpret the risk of death in parametric models and Cox model, respectively. For the statistical analysis, the statistical softwares SAS 9.1 (figures) and R (analysis) were used. In all cases, p < 0.05 was defined as the statistical significance.

One of the limitations of this study was the absence of clinical variables including type of esophageal cancer (adenocarcinoma, squamous), stage of disease, etc. The reason was absence of recorded clinical data in the Babol cancer registry and the lack of access to medical records of patients.

Univariate Gamma Frailty Model
The proportional hazards of univariate frailty model assumes that based on condition of the frailty variable U, the hazard function for the lifetime T is

\[ h_i(t / x, u) = h_0(t)u, \exp(\beta^T X) \]

where \( h_0(t) \) is the baseline hazard function with \( X = (X_1, \ldots, X_k) \) and \( \beta = (\beta_1, \ldots, \beta_k) \) as covariates and regression parameters, respectively. Here, \( u_i \) is considered a nonnegative random variable (frailty variable) for the \( i \)th (\( i=1, 2, \ldots, n \)) patient. Consequently, a frailty model is a generalization of the well-known proportional hazards model. The proportional hazards model is obtained if the frailty distribution degenerates to \( U=1 \) for all individuals. Since the hazard function cannot be negative, a positive distribution should be considered for frailty distribution. The frailty distributions most often applied are the gamma distribution, inverse Gaussian, log-normal, the positive stable distribution, Compound Poisson and a three-parameter distribution (PVF). Also, because of over-parameterization and identifiability problem, and because frailty as a random effect indicates the effect of unknown variables, it is necessary to assume that the mean of frailty equals one. We considered the frailty model with observed covariates and focused on the main aspects of univariate frailty modeling.

The density of a gamma-distributed random variable (with mean one and variance \( \sigma^2 \)) \( U \sim \Gamma(\frac{1}{\sigma^2}, \frac{1}{\sigma^2}) \) is given by

\[ f_U(u) = \frac{1}{\Gamma(\frac{1}{\sigma^2})} \left(\frac{1}{\sigma^2}\right)^{\frac{1}{\sigma^2}} \left(\frac{1}{\sigma^2}\right)^{-1} \exp\left(-\frac{u}{\sigma^2}\right) \]

Results
Of the 359 patients with esophageal cancer enrolled in this study, 225 (62.7%) were men and 134 (37.3%) were women. The mean (standard deviation) age at diagnosis was 55.23 (± 11.01) years (Table 1). The median survival time in this study reached about 9 months and 1-, 3- and 5-year survival rates following diagnosis were estimated 23%, 15% and 13%, respectively. A total of 310 (86.3%) deaths were observed (non-censored observations) during
Table 1. Characteristics of patients with esophageal cancer diagnosis

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (n=359)</th>
<th>Male (n=225)</th>
<th>Female (n=134)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place of residence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>199 (55.4)</td>
<td>134 (59.6)</td>
<td>65 (48.5)</td>
</tr>
<tr>
<td>Urban</td>
<td>160 (44.6)</td>
<td>91 (40.4)</td>
<td>69 (51.5)</td>
</tr>
<tr>
<td>Province</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mazandaran</td>
<td>188 (52.4)</td>
<td>122 (54.2)</td>
<td>66 (49.2)</td>
</tr>
<tr>
<td>Golestan</td>
<td>171 (47.6)</td>
<td>103 (45.8)</td>
<td>68 (50.8)</td>
</tr>
<tr>
<td>Family history of cancer</td>
<td>110 (30.6)</td>
<td>65 (28.9)</td>
<td>45 (33.6)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literate</td>
<td>35 (9.7)</td>
<td>29 (12.9)</td>
<td>6 (4.5)</td>
</tr>
<tr>
<td>Illiterate</td>
<td>324 (90.3)</td>
<td>196 (87.1)</td>
<td>128 (95.5)</td>
</tr>
<tr>
<td>Job</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farmer</td>
<td>186 (51.8)</td>
<td>170 (75.6)</td>
<td>16 (11.9)</td>
</tr>
<tr>
<td>Employee</td>
<td>3 (0.85)</td>
<td>3 (1.3)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Others</td>
<td>170 (47.35)</td>
<td>52 (23.1)</td>
<td>118 (88.1)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>340 (94.7)</td>
<td>217 (96.4)</td>
<td>123 (91.8)</td>
</tr>
<tr>
<td>Single</td>
<td>19 (5.3)</td>
<td>8 (3.6)</td>
<td>11 (8.2)</td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td>151 (42.1)</td>
<td>128 (56.9)</td>
<td>23 (17.1)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aryan</td>
<td>219 (61.0)</td>
<td>132 (58.7)</td>
<td>87 (64.9)</td>
</tr>
<tr>
<td>Gilak</td>
<td>11 (3.1)</td>
<td>9 (4.0)</td>
<td>2 (1.5)</td>
</tr>
<tr>
<td>Torkaman</td>
<td>92 (25.6)</td>
<td>61 (27.1)</td>
<td>31 (23.1)</td>
</tr>
<tr>
<td>Others</td>
<td>37 (10.3)</td>
<td>23 (10.2)</td>
<td>14 (10.5)</td>
</tr>
<tr>
<td>Migration status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native</td>
<td>327 (91.1)</td>
<td>203 (90.2)</td>
<td>124 (92.5)</td>
</tr>
<tr>
<td>Non-native</td>
<td>32 (8.9)</td>
<td>22 (9.8)</td>
<td>10 (7.5)</td>
</tr>
</tbody>
</table>

As the proportionality assumption of Cox model was not met in our data (p = 0.01), the Cox regression was not suitable. Even after adding frailty term (gamma) into Cox model, proportionality assumption was violated and there was no remedy in the violation of the PH assumption.

Table 2 shows the mean, median, standard deviation and confidence interval for survival time (month) of patients diagnosed with esophageal cancer, in different age groups. Based on Breslow estimator, probability value was obtained significant at 0.05 (p = 0.04); i.e., survival functions were different for different age groups (Figure 1).

Table 2. Means, Medians and Confidence Interval for Survival Time (Month) of patients with esophageal cancer, by different age groups

<table>
<thead>
<tr>
<th>Age groups</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 50</td>
<td>33.21</td>
<td>5.18</td>
<td>23.05</td>
<td>43.38</td>
<td>9.63</td>
<td>1.46</td>
<td>6.78</td>
<td>12.49</td>
</tr>
<tr>
<td>51-60</td>
<td>35.09</td>
<td>5.61</td>
<td>24.10</td>
<td>46.08</td>
<td>8.13</td>
<td>1.28</td>
<td>5.63</td>
<td>10.63</td>
</tr>
<tr>
<td>61-70</td>
<td>41.74</td>
<td>6.39</td>
<td>29.21</td>
<td>54.27</td>
<td>10.60</td>
<td>1.47</td>
<td>7.72</td>
<td>13.48</td>
</tr>
<tr>
<td>&gt; 70</td>
<td>13.39</td>
<td>3.32</td>
<td>6.88</td>
<td>19.91</td>
<td>4.70</td>
<td>0.95</td>
<td>2.83</td>
<td>6.57</td>
</tr>
<tr>
<td>Overall</td>
<td>35.10</td>
<td>3.13</td>
<td>28.97</td>
<td>41.23</td>
<td>8.97</td>
<td>0.80</td>
<td>7.39</td>
<td>10.54</td>
</tr>
</tbody>
</table>

*Estimation is limited to the largest survival time if it is censored.
Review of residual plots (Cox-Snell residuals, deviance residuals) represented a better fit of parametric models related to Cox model. Among the parametric models, the log-logistic model with gamma frailty was more appropriate to fit the data (mean deviance residuals were 0.187, 0.248 and 0.672 for log-logistic, log-normal and Cox models, respectively, Figure 2).

To enable the comparison, we calculated the AIC. Table 3 shows that according to this criterion, the log-logistic model scored best, followed by the log-normal model and the Cox model came third. According to the AIC, the log-logistic allowing for gamma heterogeneity is thus the preferred model, followed by the log-normal model. By comparing these findings, we came to the conclusion that the log-logistic model with gamma frailty is more efficient than the Cox model and log-normal model (with and without gamma frailty).

As expected, the effects of covariates were biased downwards in the Cox model when not corrected for unobserved heterogeneity in the study population (Table 4). Note that the standard deviation also increased in the gamma frailty model, and that the large standard deviation of the frailty variance \( \sigma^2 \) estimate did not exclude the possibility of no unobserved heterogeneity \( \sigma^2 = 0 \).

Table 3. Overview of the Akaike's Information Criterion Scores

<table>
<thead>
<tr>
<th>Without Heterogeneity</th>
<th>Log-likelihood</th>
<th>AIC</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>log-normal</td>
<td>-650.33</td>
<td>1326.66</td>
<td>2</td>
</tr>
<tr>
<td>log-logistic</td>
<td>-644.39</td>
<td>1314.78</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gamma Heterogeneity</th>
<th>Log-likelihood</th>
<th>AIC</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>log-normal</td>
<td>-621.48</td>
<td>1282.96</td>
<td>2</td>
</tr>
<tr>
<td>log-logistic</td>
<td>-610.69</td>
<td>1261.38</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 2. Cox-Snell residuals obtained from fitting various survival models to the esophageal cancer data.

The panels indicate the Cox-Snell residuals (together with their cumulative hazard function) obtained from fitting different parametric AFT models to the same data via maximum likelihood estimation. Obviously, the lines related to the Cox-Snell residuals of the log-normal and log-logistic models with gamma frailty are nearest to the line through the origin, indicating that these models fit the data best. In addition, the Cox model does not appear to fit the data well; it means that the proportional hazards assumption is violated. These results are consistent with our finding based on Akaike's Information Criterion.

Also, Table 4 shows that the log-normal and log-logistic models (with and without gamma frailty) have presented estimated parameters and standard error better than the Cox model. Between these two models, the log-logistic model with gamma frailty seemed more appropriate. These results were consistent with our finding based on Akaike's Information Criterion and residuals plots.

As expected, the log-logistic and log-normal analysis with gamma frailty model suggested that males faced a higher risk of the death caused by esophageal cancer than females.

Table 5 shows the results of multiple analyses for Cox models and parametric models (with and without frailty) based on hazard ratio (HR), relative risk (RR) and confidence interval for each variable. Results of multiple analyses showed a significant difference in patients who had a family history of cancer in all models.

In the log-logistic with gamma frailty model (the selected model), the relative risk reported 1.33 for family history of cancer variable indicated that the desired event (death) occurred more often (33%) in patients with positive family history of cancer compared to those without such family history.

Gender was significant under the log-normal and log-logistic with gamma frailty model but not a significant factor under the other model. This indicated that the level of the death risk due to esophageal cancer was reduced significantly for the women in the study during the follow-up period. Also, the relative risk of 1.71 for gender variable (in selected model) indicated that the event (death) occurred more often (71%) in male than in female patients.
Table 4. Parameter estimates (s.e.) in the Cox, log-normal and log-logistic models and the gamma frailty model applied to esophageal cancer data.

<table>
<thead>
<tr>
<th>variables</th>
<th>Cox parameters without frailty</th>
<th>Cox parameters with gamma frailty</th>
<th>Log-normal parameters without frailty</th>
<th>Log-normal parameters with gamma frailty</th>
<th>Log-logistic parameters without frailty</th>
<th>Log-logistic parameters with gamma frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter estimate (s.e.)</td>
<td>Parameter estimate (s.e.)</td>
<td>Parameter estimate (s.e.)</td>
<td>Parameter estimate (s.e.)</td>
<td>Parameter estimate (s.e.)</td>
<td>Parameter estimate (s.e.)</td>
</tr>
<tr>
<td>Age(years)</td>
<td>0.003 (0.005)</td>
<td>0.003 (0.007)</td>
<td>-0.004 (0.008)</td>
<td>-0.002 (0.007)</td>
<td>-0.003 (0.008)</td>
<td>-0.003 (0.006)</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>0.14 (0.17)</td>
<td>0.25 (0.21)</td>
<td>-0.34 (0.25)</td>
<td>-0.59 (0.22)*</td>
<td>-0.34 (0.24)</td>
<td>-0.53 (0.21)*</td>
</tr>
<tr>
<td>Province (Mazandaran)</td>
<td>-0.06 (0.12)</td>
<td>0.19 (0.19)</td>
<td>-0.04 (0.18)</td>
<td>-0.43 (0.19)</td>
<td>-0.06 (0.16)</td>
<td>-0.41 (0.18)</td>
</tr>
<tr>
<td>Place of residence (Urban)</td>
<td>0.06 (0.12)</td>
<td>0.03 (0.15)</td>
<td>-0.08 (0.18)</td>
<td>0.06 (0.15)</td>
<td>-0.03 (0.17)</td>
<td>0.12 (0.14)</td>
</tr>
<tr>
<td>Family history of cancer (positive)</td>
<td>0.40 (0.13)*</td>
<td>0.48 (0.16)*</td>
<td>-0.61 (0.19)*</td>
<td>-0.37 (0.16)*</td>
<td>-0.54 (0.18)*</td>
<td>-0.28 (0.15)*</td>
</tr>
<tr>
<td>Education (Literate)</td>
<td>-0.29 (0.20)</td>
<td>-0.38 (0.25)</td>
<td>0.48 (0.30)</td>
<td>0.35 (0.27)</td>
<td>0.53 (0.29)</td>
<td>0.31 (0.26)</td>
</tr>
<tr>
<td>Marital status (Married)</td>
<td>0.14 (0.27)</td>
<td>0.18 (0.33)</td>
<td>-0.19 (0.40)</td>
<td>-0.07 (0.35)</td>
<td>-0.17 (0.39)</td>
<td>-0.13 (0.31)</td>
</tr>
<tr>
<td>Cigarette smoking (smoking)</td>
<td>0.16 (0.13)</td>
<td>0.17 (0.16)</td>
<td>-0.16 (0.20)</td>
<td>0.07 (0.18)</td>
<td>-0.12 (0.19)</td>
<td>0.02 (0.16)</td>
</tr>
<tr>
<td>Migration status (Native)</td>
<td>-0.18 (0.21)</td>
<td>-0.07 (0.27)</td>
<td>0.11 (0.31)</td>
<td>-0.17 (0.28)</td>
<td>0.14 (0.31)</td>
<td>-0.13 (0.27)</td>
</tr>
<tr>
<td>Job (Employee)</td>
<td>-0.53 (0.72)</td>
<td>-0.73 (0.83)</td>
<td>1.04 (0.99)</td>
<td>1.04 (0.85)</td>
<td>0.93 (0.99)</td>
<td>0.78 (0.77)</td>
</tr>
<tr>
<td>Job (Others)</td>
<td>-0.02 (0.15)</td>
<td>0.01 (0.19)</td>
<td>-0.07 (0.22)</td>
<td>-0.11 (0.19)</td>
<td>-0.04 (0.21)</td>
<td>-0.09 (0.18)</td>
</tr>
<tr>
<td>Ethnicity (Gilak)</td>
<td>-0.11 (0.33)</td>
<td>-0.30 (0.40)</td>
<td>0.49 (0.51)</td>
<td>0.93 (0.43)</td>
<td>0.54 (0.47)</td>
<td>0.81 (0.39)</td>
</tr>
<tr>
<td>Ethnicity (Torkaman)</td>
<td>0.34 (0.18)</td>
<td>0.42 (0.22)</td>
<td>-0.42 (0.27)</td>
<td>-0.23 (0.23)</td>
<td>-0.44 (0.25)</td>
<td>-0.32 (0.20)</td>
</tr>
<tr>
<td>Ethnicity (Others)</td>
<td>0.32 (0.20)</td>
<td>0.43 (0.25)</td>
<td>-0.38 (0.29)</td>
<td>-0.28 (0.25)</td>
<td>-0.43 (0.28)</td>
<td>-0.41 (0.23)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>---</td>
<td>1.38 (0.84)</td>
<td>---</td>
<td>0.82 (0.14)</td>
<td>---</td>
<td>0.76 (0.13)</td>
</tr>
</tbody>
</table>

* Significant at 0.05 level

Neither parametric models nor Cox model showed age, place of residence, province, educational level, smoking, job, marital status, ethnicity and migration status as prognostic factors.

Discussion

Esophageal cancer is one of the most common cancers in Iran.27 The cancer is a particularly devastating form of cancer with a relatively low survival rate, and people generally will not live a long time after diagnosis. The five-year survival rate in this study was 13% which is lower than that of many other countries.45-48 This may be accounted for by the fact that Iranian patients generally seek medical advice and delayed diagnosis, when the disease has reached an advanced stage. Several factors are known in various studies as influencing prognosis factors and have been introduced.49-58

In the literature, there are many studies on the field of cancer, but researchers tend to examine the effects of covariates on patients survival using Cox regression model instead of parametric ones. A systematic study on cancer journals showed that only in 5% of cancer studies in which Cox regression model is used, the assumptions of the model have been investigated.59 If presumptions are not met, results of Cox model are seriously under question. As an alternative, parametric models such as log-normal, log logistic, Weibull and exponential can be employed. The only assumption of parametric models is that the variable time follows a specific distribution.30,37

This study assessed the relationship between survival of patients with esophageal cancer and numerous prognostic factors such as age at diagnosis, sex, family history of cancer, marital status, smoking status, ethnicity,
medication status, education, place of residence and migration status.

Gender was a strong and independent prognostic factor, and our finding in multiple analyses was similar to the previous reports indicating better survival for female patients. This means that women getting affected by esophageal cancer have a slightly longer survival than men in Northern Iran; these findings are consistent with those obtained by European countries.\textsuperscript{1,60-62}

In this study, family history of cancer is another important prognostic factor of esophageal cancer and many studies indicate that the survival depends on the presence of family history of cancer. These findings are aligned with the results of many studies in this field.\textsuperscript{63} Yue-quan et al. showed that the five prognostic factors determined as significant by $p$ value were tumor size, grade of differentiation, lymphadenopathy, stage of cancer, and family history of esophageal cancer after esophagectomy.

Family history of esophageal cancer is an important prognostic factor that surgeons should take into consideration when selecting a treatment method.\textsuperscript{64}

Statistical assessment of considered models using AIC criteria revealed the log-logistic model with gamma frailty to be the most powerful one in predicting survival of cancerous patients when compared to Cox and log-normal models.
There are preferences to use parametric models, because of good discrimination, provided censoring percentage not to exceed 40-50 percent; the condition our data was satisfied by censoring rate of close to 14 percent, so results of parametric models were considered acceptable.

In Nardi et al. study, Cox model and alternative parametric models in three clinical studies were compared. They used normal-deviate residuals for evaluation of parametric models assumptions. Nardi et al. also studied Weibull model based on the estimated variation of parameter rate criteria and showed that it was better than the other models. In our study, the same result was obtained by log-logistic model with gamma frailty. Orbe et al. in a simulation study, compared Cox regression and accelerated failure time (AFT) models. They used the proposed method by Stute for fitting linear regression models with right-censored data. The results showed that if the proportional hazards assumption is violated or this assumption is met, log logistic, log-normal and Stute models are more efficient than the Cox model.

Bradburn et al. evaluated the adequacy type of parametric models and Cox proportional hazard model through residuals and AIC criterion. In their study performed on patients with ovarian and lung cancer, generalized gamma model compared to Cox and other parametric models reached a higher log-likelihood and lower AIC that make it more efficient model.

In Cox and parametric models, hazard function may depend on unknown or non-measurable factors which can cause the regression coefficients estimated from such models to be biased. In consequence, in order to overcome the problem and better model survival of patients, the frailty models were introduced. In fact, these models are used to explain the random variation of survival function due to unknown risk factors, such as genetic factors and numerous environmental factors. Vaupel et al. for the first time proposed frailty in order to describe the consequences of the existence of multiple variation source for univariate lifetime data.

Random effects models are called frailty models in survival analysis. These models are relatively new in survival and widely studied in the 1990's and now are being used as the subject of many investigations. Technical problems in estimating the parameters have been caused the Cox model to be used less. Henderson and Oman in a theoretical method revealed that in case of non-use of frailty model when there is frailty effect, bias may occur in the estimates of regression coefficients. Schumacher et al. showed how to delete an important factor and reduce the relative risk estimates. Keiding et al. showed that removing one of the two explanatory variables might increase the hazard variance function and cause bias in estimating the other variable in the model. They also suggested that in order to account for the effect of unknown variables, in univariate survival data, it might be better to use the accelerated failure time (AFT) models.

**Conclusion**

In conclusion, the overall survival rates for patients with esophageal cancer in Northern Iran are relatively low. Since stage of cancer is a very important factor influencing survival of patients with esophageal cancer, these short survival rates can be due to the fact that patients with esophageal cancer in Northern Iran are, in general, referred to physicians at late stages of the disease.

Our study indicated that gender and family history of cancer were related factors to lifetime survival in patients with esophageal cancer. According to our findings, the early recognition of family history of cancer and, in consequence, awareness of family members to consider the possibility of family screening may result in a decrease in death rate due to esophageal cancer. Steady public and professional education is required to increase awareness of hereditary esophageal cancer and the
possibility of family screening, early diagnosis and therapy. Also, we recommend adding psychosocial support for these at-risk patients and their families as well as preventive lifestyle and dietary intervention. Furthermore, women experience reduced risk of death due to esophageal cancer than men. Also, comparing parametric models and Cox model showed that log-logistic with gamma frailty model (AIC and mean deviance residual has the lowest values compared to other models) can be a useful statistical model to find prognostic factors.

**Acknowledgements**
Authors are urged to express their gratitude to the National Institute of Health Research (NIHR) and Tehran University of Medical Sciences for their support in data collection, financial support and collaboration in this study.

**Conflict of Interests**
Authors have no conflict of interests.

**Authors' Contributions**
All the authors have carried out the study, participated in the design of the study and acquisition of data performed the statistical analysis and wrote the manuscript. All authors read and approved the final manuscript.

**References**


